

BEST AVAILABLE COPY

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日
Date of Application: 2 0 0 3 年 9 月 1 日

出 願 番 号
Application Number: 特 願 2 0 0 3 - 3 0 9 1 9 4
[ST. 10/C]: [J P 2 0 0 3 - 3 0 9 1 9 4]

出 願 人
Applicant(s): 株式会社日立製作所

CERTIFIED COPY OF
PRIORITY DOCUMENT

2 0 0 4 年 3 月 2 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康 夫

出証番号 出証特 2 0 0 4 - 3 0 1 5 5 9 6

【書類名】 特許願
【整理番号】 HI030480
【提出日】 平成15年 9月 1日
【あて先】 特許庁長官殿
【国際特許分類】 G06F 3/06
【発明者】
 【住所又は居所】 神奈川県小田原市中里 3 2 2 番 2 号 株式会社日立製作所 R A I
 Dシステム事業部内
 【氏名】 高橋 正充
【発明者】
 【住所又は居所】 神奈川県小田原市中里 3 2 2 番 2 号 株式会社日立製作所 R A I
 Dシステム事業部内
 【氏名】 佐藤 孝夫
【発明者】
 【住所又は居所】 神奈川県小田原市中里 3 2 2 番 2 号 株式会社日立製作所 R A I
 Dシステム事業部内
 【氏名】 小澤 匡二
【特許出願人】
 【識別番号】 000005108
 【氏名又は名称】 株式会社日立製作所
【代理人】
 【識別番号】 110000176
 【氏名又は名称】 一色国際特許業務法人
 【代表者】 一色 健輔
【手数料の表示】
 【予納台帳番号】 211868
 【納付金額】 21,000円
【提出物件の目録】
 【物件名】 特許請求の範囲 1
 【物件名】 明細書 1
 【物件名】 図面 1
 【物件名】 要約書 1

【書類名】 特許請求の範囲**【請求項 1】**

第一のサイトに設置されている第一の記憶ボリュームを備える第一の記憶装置と、
第二のサイトに設置されている第二の記憶ボリュームを備える第二の記憶装置と、
第三のサイトに設置されている第三の記憶ボリュームを備える第三の記憶装置と
が互いに接続可能に構成されるストレージシステムの制御方法において、
第一の時刻において前記第一の記憶ボリュームに記憶されているデータの複製を前記第二の記憶ボリュームに記憶するステップと、
前記第一の記憶ボリュームに書き込まれたデータの複製を前記第三の記憶ボリュームに書き込むステップと、
前記第一の時刻以降に前記第一の記憶ボリュームに書き込まれたデータの書き込み履歴を、前記第三の記憶装置に第一の差分管理テーブルとして記憶するステップと、
前記第三の記憶装置が、前記第三の記憶装置の前記第一の差分管理テーブルと前記第三の記憶ボリュームとを用いて、前記第二の記憶ボリュームに記憶されているデータの内容を前記第一の記憶ボリュームに記憶されているデータの内容に一致させるステップとを備えることを特徴とするストレージシステムの制御方法。

【請求項 2】

請求項 1 に記載のストレージシステムの制御方法において、
前記第一の時刻以降に前記第一の記憶ボリュームに書き込まれたデータの書き込み履歴を前記第一の差分管理テーブルとして記憶する前記ステップは、
前記第三の記憶装置が、前記第一の記憶装置から受信したデータの書き込み要求に設定されている時刻と前記第一の時刻とを比較するステップと、
前記第三の記憶装置が、前記第一の記憶装置から受信した前記データの前記書き込み要求に設定されている前記書き込み時刻が前記第一の時刻以降である場合には、前記データの書き込み履歴を前記第一の差分管理テーブルに記憶するステップとを備えることを特徴とするストレージシステムの制御方法。

【請求項 3】

請求項 1 に記載のストレージシステムの制御方法において、
前記第三の記憶装置が、前記第一の時刻以降に前記第一の記憶ボリュームに書き込まれたデータの書き込み履歴を記録した前記第一の差分管理テーブルと、前記第一の時刻以降の第二の時刻以降に前記第一の記憶ボリュームに書き込まれたデータの書き込み履歴を記録した第二の差分管理テーブルとを記憶するステップを備えることを特徴とするストレージシステムの制御方法。

【請求項 4】

第一のサイトに設置されている第一の記憶ボリュームを備える第一の記憶装置と、
第二のサイトに設置されている第二の記憶ボリュームを備える第二の記憶装置と、
第三のサイトに設置されている第三の記憶ボリュームを備える第三の記憶装置と
が互いに接続可能に構成されるストレージシステムにおいて、
第一の時刻において前記第一の記憶ボリュームに記憶されているデータの複製を前記第二の記憶ボリュームに記憶する手段と、
前記第一の記憶ボリュームに書き込まれたデータの複製を前記第三の記憶ボリュームに書き込む手段と、
前記第一の時刻以降に前記第一の記憶ボリュームに書き込まれたデータの書き込み履歴を、前記第三の記憶装置に第一の差分管理テーブルとして記憶する手段と、
前記第三の記憶装置が、前記第三の記憶装置の前記第一の差分管理テーブルと前記第三の記憶ボリュームとを用いて、前記第二の記憶ボリュームに記憶されているデータの内容を前記第一の記憶ボリュームに記憶されているデータの内容に一致させる手段とを備えることを特徴とするストレージシステム。

【請求項 5】

請求項 4 に記載のストレージシステムにおいて、

前記第一の時刻以降に前記第一の記憶ボリュームに書き込まれたデータの書き込み履歴を前記第一の差分管理テーブルとして記憶する前記手段は、

前記第三の記憶装置が、前記第一の記憶装置から受信したデータの書き込み要求に設定されている時刻と前記第一の時刻とを比較する手段と、

前記第三の記憶装置が、前記第一の記憶装置から受信した前記データの前記書き込み要求に設定されている前記書き込み時刻が前記第一の時刻以降である場合には、前記データの書き込み履歴を前記第一の差分管理テーブルに記憶する手段とを備えることを特徴とするストレージシステム。

【請求項 6】

請求項 4 に記載のストレージシステムにおいて、

前記第三の記憶装置が、前記第一の時刻以降に前記第一の記憶ボリュームに書き込まれたデータの書き込み履歴を記録した前記第一の差分管理テーブルと、前記第一の時刻以降の第二の時刻以降に前記第一の記憶ボリュームに書き込まれたデータの書き込み履歴を記録した第二の差分管理テーブルとを記憶する手段を備えることを特徴とするストレージシステム。

【請求項 7】

データを記憶するための記憶ボリュームと、第一の時刻以降に前記記憶ボリュームに書き込まれたデータの書き込み履歴を第一の差分管理テーブルとして記憶する手段と、

前記第一の時刻以降の第二の時刻以降に前記記憶ボリュームに書き込まれたデータの書き込み履歴を第二の差分管理テーブルとして記憶する手段とを備えることを特徴とするストレージ装置。

【書類名】明細書

【発明の名称】 ストレージシステムの制御方法、ストレージシステム、及びストレージ装置

【技術分野】**【0001】**

本発明はストレージシステムの制御方法、及びストレージシステムに関する。

【背景技術】**【0002】**

情報処理システムにおける災害復旧（ディザスタリカバリ）が注目されている。ディザスタリカバリを実現する技術として、プライマリサイトに設置されている記憶装置のデータの複製を、これとは遠隔したリモートサイトに設置されている記憶装置においても管理する技術（以下、この技術を「リモートコピー」と称する）が知られている。プライマリサイトの被災時に、リモートサイトに設置されている記憶装置のデータを使用することで、プライマリサイトで行われていた処理をリモートサイトで継続して行うことができる。

【発明の開示】**【発明が解決しようとする課題】****【0003】**

前述の方法において、プライマリサイトの被災時にプライマリサイトで行われていた処理をリモートサイトで継続して行うためには、プライマリサイトからリモートサイトへのリモートコピーをリアルタイムに行う必要があった。しかしながら、プライマリサイトとリモートサイトとの間は遠距離である場合が多く、プライマリサイトからリモートサイトへのリモートコピーをリアルタイムに行うと、データの通信に時間がかかり、プライマリサイトの処理性能が低下していた。高可用性（HA: High Availability）が求められる情報処理システムにおいては、プライマリサイトの処理性能を低下させず、さらに、プライマリサイトが被災した際に、リモートサイトで迅速に処理を再開させることが求められていた。

【0004】

本発明は上記課題を鑑みてなされたものであり、ストレージシステムの制御方法、及びストレージシステムを提供することを目的とする。

【課題を解決するための手段】**【0005】**

上記課題を解決するために、本発明に係るストレージシステムは、第一のサイトに設置された第一の記憶ボリュームを備える第一の記憶装置と、

第二のサイトに設置された第二の記憶ボリュームを備える第二の記憶装置と、

第三のサイトに設置された第三の記憶ボリュームを備える第三の記憶装置とが互いに接続可能に構成されるストレージシステムの制御方法において、

前記第一の記憶ボリュームの第一の時刻における複製を前記第二の記憶ボリュームに記憶するステップと、

前記第一の記憶ボリュームに書き込まれたデータの複製を前記第三の記憶ボリュームに書き込むステップと、

前記第一の時刻以降に前記第一の記憶ボリュームに書き込まれたデータの書き込み履歴を、前記第三の記憶装置に第一の差分管理テーブルとして記憶するステップと、

前記第二の記憶装置が、前記第三の記憶装置の前記第一の差分管理テーブルと前記第三の記憶ボリュームを用いて、前記第二の記憶ボリュームを前記第一の記憶ボリュームと一致させるステップとを備える。

【0006】

ここで、第一のサイトをプライマリサイト、第二のサイトをリモートサイト、第三のサイトを第一のサイトの近隣に設けたローカルサイトとして運用する情報処理システムにおいて、第一のサイトが被災した場合に、第一の記憶ボリュームと第二の記憶ボリュームの差分データのみを、第三の記憶ボリュームから第二の記憶ボリュームにリモートコピーす

ることで、第二の記憶ボリュームを第一の記憶ボリュームと一致させることができるようになる。

【発明の効果】

【0007】

ストレージシステムの制御方法、及びストレージシステムを提供することができる。

【発明を実施するための最良の形態】

【0008】

===情報処理システムの構成===

図1に本実施の形態に係るストレージシステム100を含む情報処理システムの全体構造を示すブロック図を示す。

【0009】

情報処理システムは、第一のサイトに設置されている記憶装置10（以後、「第一の記憶装置」と称する）と、これにアクセスする情報処理装置11（以後、「第一の情報処理装置」と称する）と、第二のサイトに設置されている記憶装置20（以後、「第二の記憶装置」と称する）と、これにアクセスする情報処理装置21（以後、「第二の情報処理装置」と称する）と、第三のサイトに設置されている記憶装置30（以後、「第三の記憶装置」と称する）とを含んで構成される。各サイトは、具体的には、例えば、大学や企業などの組織によって運営されるコンピュータ施設や、インターネット上のWebサーバの運営、ASP（Application Service Provider）の運用などが行われているデータセンタなどである。この情報処理システムは、地震・火災・台風・洪水・落雷・テロなどに対するディザスタリカバリの実現のために構築される。

【0010】

本実施の形態においては、第一のサイトが前述のプライマリサイト、第二のサイトが前述のリモートサイト、第三のサイトが前述のローカルサイトの構成となっている。

【0011】

各サイトに設置されている記憶装置10、20、30は第一のネットワーク40を介して通信可能に接続されている。第一のネットワーク40は、例えば、ギガビットイーサネット（登録商標）、ATM（Asynchronous Transfer Mode）、公衆回線などである。

【0012】

各情報処理装置11、21は、CPU（Central Processing Unit）やメモリを備えるコンピュータであり、パーソナルコンピュータやワークステーション、メインフレームなどのコンピュータである。各情報処理装置11、21は、結合された複数台のコンピュータで構成されることもある。各情報処理装置11、21ではオペレーティングシステムが動作している。オペレーティングシステム上ではアプリケーションソフトウェアが動作している。

【0013】

アプリケーションソフトウェアは、例えば、銀行の自動預金預け払いシステムや航空機の座席予約システムの機能を提供する。第一のサイト及び第二のサイトにおける情報処理装置11、21と記憶装置10、20とはそれぞれ通信手段（通信線、ネットワーク）により接続されている。通信手段は、例えば、LAN（Local Area Network）やSAN（Storage Area Network）、iSCSI（Internet Small Computer System Interface）、Fibre Channel、ESCON（Enterprise Systems Connection）（登録商標）、FICON（Fibre Connection）（登録商標）などである。

【0014】

図2に第一乃至第三の記憶装置10、20、30の一例として説明するディスクアレイ装置の具体的な構成を示している。なお、第一乃至第三の記憶装置10、20、30は、ディスクアレイ装置以外にも、例えば、半導体記憶装置などであってもよい。ディスクアレイ装置は、チャンネル制御部101、リモート通信インタフェース102、ディスク制御部103、共有メモリ104、キャッシュメモリ105、これらの間を通信可能に接続するクロスバススイッチなどで構成されるスイッチング制御部106、管理端末107、及び

記憶デバイス 108などを備えて構成される。

【0015】

キャッシュメモリ 105は、主としてチャネル制御部 101とディスク制御部 103との間で授受されるデータを一時的に記憶するために用いられる。例えばチャネル制御部 101が情報処理装置 11, 21から受信したデータ入出力コマンドが書き込みコマンドである場合には、チャネル制御部 101は情報処理装置 11, 21から受信した書き込みデータをキャッシュメモリ 105に書き込む。またディスク制御部 103はキャッシュメモリ 105から書き込みデータを読み出して記憶デバイス 108に書き込む。

【0016】

ディスク制御部 103は、チャネル制御部 101により共有メモリ 104に書き込まれたデータ I/O要求を読み出してそのデータ I/O要求に設定されているコマンド（例えば、SCSI (Small Computer System Interface) 規格のコマンド）に従って記憶デバイス 108にデータの書き込みや読み出しなどの処理を実行する。ディスク制御部 103は記憶デバイス 108から読み出したデータをキャッシュメモリ 105に書き込む。またデータの書き込み完了通知や読み出し完了通知などをチャネル制御部 101に送信する。ディスク制御部 103は、記憶デバイス 108をいわゆる RAID (Redundant Array of Inexpensive Disks) 方式に規定される RAID レベル（例えば、0, 1, 5）で制御する機能を備えることもある。

【0017】

記憶デバイス 108は、例えば、ハードディスク装置である。記憶デバイス 108はディスクアレイ装置と一体型とすることもできるし、別体とすることもできる。各サイトの記憶デバイス 108により提供される記憶領域は、この記憶領域上に論理的に設定されるボリュームである論理ボリューム 109を単位として管理されている。記憶デバイス 108へのデータの書き込みや読み出しは、論理ボリューム 109に付与される識別子を指定して行なうことができる。

【0018】

管理端末 107はディスクアレイ装置や記憶デバイス 108を保守・管理するためのコンピュータである。チャネル制御部 101やディスク制御部 103において実行されるソフトウェアやパラメータの変更は、管理端末 107からの指示により行われる。管理端末 107はディスクアレイ装置に内蔵される形態とすることもできるし、別体とすることもできる。

【0019】

リモート通信インタフェース 102は、他の記憶装置 10, 20, 30とデータ伝送をするための通信インタフェース（チャネルエクステンダ）であり、後述するリモートコピーにおける複製データの伝送はこのリモート通信インタフェース 102を介して行われる。リモート通信インタフェース 102は、チャネル制御部 101のインタフェース（例えば、Fibre Channel、ESCON（登録商標）、FICON（登録商標）などのインタフェース）を第一のネットワーク 40の通信方式に変換する。これにより他の記憶装置 10, 20, 30との間でのデータ伝送が実現される。

【0020】

なお、ディスクアレイ装置は、以上に説明した構成のもの以外にも、例えば、NFS (Network File System) などのプロトコルにより情報処理装置 11, 21からファイル名指定によるデータ入出力要求を受け付けるように構成されたNAS (Network Attached Storage) として機能するものなどであってもよい。

【0021】

共有メモリ 104はチャネル制御部 101とディスク制御部 103の両方からアクセスが可能である。データ入出力要求コマンドの受け渡しに利用される他、記憶装置 10, 20, 30や記憶デバイス 108の管理情報等が記憶される。本実施の形態においては、図 5に示すようにコンシステンシグループ管理テーブル 200、ペア管理テーブル 201、差分管理テーブル 202が共有メモリ 104に記憶される。

【0022】

情報処理装置 11, 21 の構成を示すブロック図を図 3 に示す。

情報処理装置 11, 21 は、CPU 110、メモリ 120、ポート 130、記録媒体読取装置 140、入力装置 150、出力装置 160 を備える。

【0023】

CPU 110 は情報処理装置 11, 21 の全体の制御を司るもので、メモリ 120 に格納されたプログラムを実行することにより各種機能を実現する。記録媒体読取装置 140 は、記録媒体 170 に記録されているプログラムやデータを読み取るための装置である。読み取られたプログラムやデータはメモリ 120 に格納される。従って、例えば記録媒体 170 に記録された記憶デバイス管理プログラム 121 やアプリケーションプログラム 122 を、記録媒体読取装置 140 を用いて上記記録媒体 170 から読み取って、メモリ 120 に格納するようにすることができる。記録媒体 170 としてはフレキシブルディスクや CD-ROM、半導体メモリ等を用いることができる。記録媒体読取装置 140 は情報処理装置 11, 21 に内蔵されている形態とすることもできるし、外付されている形態とすることもできる。入力装置 150 はオペレータ等による情報処理装置 11, 21 へのデータ入力等のために用いられる。入力装置 150 としては例えばキーボードやマウス等が用いられる。出力装置 160 は情報を外部に出力するための装置である。出力装置 160 としては例えばディスプレイやプリンタ等が用いられる。ポート 130 は記憶装置 10, 20, 30 と通信を行うための装置である。また他の情報処理装置 11, 21 との間で通信を行うために使用することもできる。この場合、例えば記憶デバイス管理プログラム 121 やアプリケーションプログラム 122 を、ポート 130 を介して他の情報処理装置 11, 21 から受信して、メモリ 120 に格納するようにすることもできる。

【0024】

記憶デバイス管理プログラム 121 は、記憶デバイス 108 が備える論理ボリューム 109 に記憶されるデータの複製管理を行うためのプログラムである。複製管理を行うための各種コマンドを記憶装置 10, 20, 30 に対して送信することにより複製管理が行われる。

【0025】

アプリケーションプログラム 122 は、情報処理装置 11, 21 により各種機能を実現するためのプログラムである。例えば上述した銀行の自動預金預け払いシステムを実現するための機能や航空機の座席予約システムを実現するための機能等である。

【0026】

チャネル制御部 101 の構成を示すブロック図を図 4 に示す。

チャネル制御部 101 は、CPU 211、キャッシュメモリ 212、制御メモリ 213、ポート 215、バス 216 を備えている。

【0027】

CPU 211 はチャネル制御部 101 の全体の制御を司るものであり、制御メモリ 213 に格納された制御プログラム 214 を実行する。制御メモリ 213 に格納された制御プログラム 214 が実行されることにより本実施の形態に係るデータの複製管理が実現される。キャッシュメモリ 212 は情報処理装置 11, 21 との間で授受されるデータやコマンド等を一時的に格納するためのメモリである。ポート 215 は情報処理装置 11, 21 との間の通信や記憶装置 10, 20, 30 内部の他の装置との間の通信を行うための通信インタフェースである。バス 216 はこれらの装置を相互に接続する。

【0028】

====ペア管理テーブル====

ペア管理テーブル 201 は、各記憶装置 10, 20, 30 の論理ボリューム 109 の複製管理を行うためのテーブルである。ペア管理テーブル 201 は、「ペア種類」、「複製方式」欄、「複製元装置」欄、「複製先装置」欄、「複製元ボリューム」欄、「複製先ボリューム」欄、「ペア状態」欄、「コンシステンシグループ」欄を備える。

【0029】

ペアとは2つの論理ボリューム109により形成される論理ボリューム109の組み合わせを言う。また、ペアを形成する2つの論理ボリューム109が同一の記憶装置10, 20, 30にある場合を「ローカルペア」、異なる記憶装置10, 20, 30にある場合を「リモートペア」と表現する。ペアを形成する論理ボリューム109は、一方を主論理ボリューム、他方を副論理ボリュームとして管理される。一つの主論理ボリュームに対して複数の副論理ボリュームを組み合わせることも可能である。

【0030】

情報処理装置11, 21が複製元の記憶装置10, 20, 30にリモートペアの形成を指示すると、複製元の記憶装置10, 20, 30は複製元の記憶装置10, 20, 30内の共有メモリ104のペア管理テーブル202を更新する。その後、複製元の記憶装置10, 20, 30は、複製先の記憶装置10, 20, 30に当該リモートペアの形成を指示する。これを受信した複製先の記憶装置10, 20, 30は、複製先の記憶装置10, 20, 30内の共有メモリ104のペア管理テーブル202を更新する。

【0031】

ペア管理テーブル201の「ペア種類」欄は当該ペアがローカルペアであるかリモートペアであるかを示す。「複製方式」欄は、当該ペアがリモートペアである場合に、リモートコピーの方式が同期方式であるか非同期方式であるかを示す。なお、リモートコピー及びその方式については後述する。「複製元装置」欄と「複製先装置」欄は、当該ペアがリモートペアである場合に、複製元の記憶装置10, 20, 30と複製先の記憶装置10, 20, 30を示す。「複製元ボリューム」欄は当該ペアの主論理ボリュームを、「複製先ボリューム」欄は当該ペアの副論理ボリュームを示す。

【0032】

「ペア状態」欄は当該ペアの状態を示す。ペアの状態としては、「ペア中」、「スプリット中」、及び「リシンク中」がある。

「ペア中」の場合は、情報処理装置11, 21から主論理ボリュームに書き込まれたデータが副論理ボリュームにも反映される。このような主論理ボリュームと副論理ボリュームとの対応付けにより、主論理ボリュームに記憶されている内容と副論理ボリュームに記憶されている内容との同一性を確保することができる。

「スプリット中」の場合は、情報処理装置11, 21から主論理ボリュームにデータが書き込まれても副論理ボリュームには反映されない。すなわち「スプリット中」の状態は、「ペア中」における主論理ボリュームと副論理ボリュームとの対応付けが解除された状態である。従って主論理ボリュームの内容と副論理ボリュームの内容の同一性は保証されない。しかし「スプリット中」は副論理ボリュームの内容が更新されないため、この間にデータのバックアップをとることが可能となる。

「リシンク中」は「スプリット中」から「ペア中」に移行する途中の状態である。すなわち、「スプリット中」に主論理ボリュームに対してなされたデータの更新を副論理ボリュームに反映している状態である。反映が完了すると当該ペアの状態は「ペア中」になる。なお、「スプリット中」のペアを「ペア中」の状態に移行することを、ペアを再形成すると言う。

【0033】

上記ペアの形成、ペアのスプリット、及びペアの再形成は、記憶デバイス管理プログラム121が実行されている情報処理装置11, 21に対して入力装置150からオペレータが指示入力を与えることにより行われる。オペレータにより与えられた指示入力は記憶装置10, 20, 30のチャンネル制御部101に送信される。チャンネル制御部101は制御プログラム214を実行して、上記指示に従って、ペアの形成やペア状態の変更を行う。チャンネル制御部101は形成したペアのペア状態に応じて、例えば「ペア中」のペアに対しては主論理ボリュームの更新データの複製を副論理ボリュームへ反映する等、論理ボリューム109に対する制御を行う。

【0034】

チャンネル制御部101によるペア状態の変更は、各ペアに対して逐次に行われる。なぜ

ならば、例えば前述したように一つの主論理ボリュームに対して複数の副論理ボリュームを組み合わせてペアを形成することもできるので、かかる場合に同時に複数のペア状態が変化すると、主論理ボリュームの管理が複雑になるからである。

【0035】

なお、上記ペアの形成やペア状態の移行は、情報処理装置 11, 21 がオペレータから受けた指示により開始される他、オペレータからの指示がなくても、例えば所定の時刻になると自動的にペア状態の移行が行われるようにすることもできる。またポート 215 を介して接続される他の情報処理装置 11, 21 から指示を受信することにより行われるようにすることもできる。

【0036】

本実施の形態においては、ペアは図 6 のペア管理テーブルのように設定されている。つまり、第一のサイトの論理ボリューム 109（以後、「第一の論理ボリューム」と称する）がローカルペアを形成している。また、第一の論理ボリュームのローカルペアの副論理ボリューム（以後、「第一の副論理ボリューム」と称する）と第二のサイトの論理ボリューム 109（以後、「第二の論理ボリューム」と称する）がリモートペアを形成している。第一の主論理ボリュームと第三のサイトの論理ボリューム 109（以後、「第三の論理ボリューム」と称する）とのリモートペアは常に「ペア中」となっており、後述する同期方式でのリモートコピーにより、第三の論理ボリュームは常に第一の主論理ボリュームと一致した状態となっている。

【0037】

また、第一の論理ボリュームから第二の論理ボリュームへのデータのバックアップは次のように行われている。まず、第一の情報処理装置 11 からの指示により、第一の記憶装置 10 が第一の論理ボリュームのローカルペアを「スプリット中」に移行する。スプリットが完了すると、第一の情報処理装置 11 からの指示により、第一の記憶装置 10 が第一の副論理ボリュームと第二の論理ボリュームのリモートペアを再形成する。なお、当該リモートペアの再形成中も、第一の情報処理装置は第一の論理ボリュームのローカルペアの主論理ボリューム（以後、「第一の主論理ボリューム」と称する）を用いて処理を継続することができる。

【0038】

===コンシステンシグループ===

ペア状態管理テーブル 201 の「コンシステンシグループ」欄には、当該ペアが属するコンシステンシグループ（ペアグループ）の番号が記載される。ここでコンシステンシグループとは、スプリット状態への移行が同時に行われるように制御される複数の論理ボリューム 109 のペアにより構成されるグループをいう。つまり、前述のように複数のペアに対するペア状態の変更は各ペアに対して逐次に行われるが、同一コンシステンシグループに属する複数のペアに対しては、スプリット状態への移行が同時に行われるように制御される（以下、スプリット状態への移行の同時性という）。

【0039】

例えば、記憶装置 10, 20, 30 が同一コンシステンシグループに属する複数のペアを順次「ペア中」から「スプリット中」に変更している間に、情報処理装置 11, 21 から記憶装置 10, 20, 30 へデータの書き込みがあった場合を考える。コンシステンシグループが形成されていなければ、当該データの書き込みがスプリット状態への移行が完了したペアの主論理ボリュームに対するものであった場合には当該書き込みデータは副論理ボリュームには反映されない。逆に、スプリット状態への移行が完了していないペアの主論理ボリュームに対するものであった場合には当該書き込みデータは副論理ボリュームに反映されてしまう。しかしながら、コンシステンシグループが形成されている場合には、スプリット状態への移行が完了している、していないに拘わらず、当該書き込みデータは副論理ボリュームへは反映されない。なぜならば、情報処理装置 11, 21 から記憶装置 10, 20, 30 への当該データの書き込みは、コンシステンシグループのペアに対するスプリットを開始（対応付けの解除を開始）した後に発行されたものだからである。

【0040】

このように複数のペアに対してコンシステンシグループを形成することは、特に、例えば書き込みデータが一つの論理ボリューム109に収まりきらないほど大きな場合や、一つのファイルデータを複数の論理ボリューム109に格納するように制御している場合等、一つのデータが複数の論理ボリューム109に跨って格納されている場合に有効である。

【0041】

なお、コンシステンシグループによる各ペアのスプリット状態への移行の同時性の確保は、情報処理装置11, 21から副論理ボリュームへのデータの書き込み要求又は読み出し要求があった場合にも有効である。

すなわち、仮にコンシステンシグループが形成されていなければ、スプリット状態への移行が完了したペアの副論理ボリュームに対してはデータの書き込み又は読み出しが可能であるのに対し、スプリット状態への移行が完了していないペアの副論理ボリュームに対してはデータの書き込み又は読み出しが禁止されているからである。

また、コンシステンシグループのペアに対するスプリットは、スプリット開始時刻を指定して行うことができる。スプリットの開始時刻は、情報処理装置11, 21から送信されるコマンドにより指示される。

【0042】

====リモートコピー====

第一乃至第三の記憶装置10, 20, 30の論理ボリューム109で前述のリモートペアが形成されている場合は、複製元の記憶装置10, 20, 30が複製先の記憶装置10, 20, 30に、リモートコピーを用いて複製したデータの伝送を行っている。複製元の記憶装置10, 20, 30の主論理ボリュームと複製先の記憶装置10, 20, 30の副論理ボリュームとの対応付けは、前述したペア管理テーブル201に設定されている。ペア状態が「ペア中」の間に情報処理装置11, 21が複製元の記憶装置10, 20, 30の主論理ボリュームにデータを書き込むと、複製元の記憶装置10, 20, 30は、そのデータの複製を第一のネットワーク40を介して複製先の記憶装置10, 20, 30に送信する。複製先の記憶装置10, 20, 30は、受信したデータを副論理ボリュームに書き込む。ペア状態が「スプリット中」の間は、情報処理装置11, 21が複製元の記憶装置10, 20, 30の主論理ボリュームにデータを書き込んでも、複製元の記憶装置10, 20, 30は複製先の記憶装置10, 20, 30に当該データの複製を送信しない。複製元の記憶装置10, 20, 30は、ペア状態が「リシンク中」に移行されたタイミングで複製先の記憶装置10, 20, 30に当該データの複製を送信し、複製先の記憶装置10, 20, 30は、受信したデータを副論理ボリュームに書き込む。すなわち、リモートコピーは、対応付けられている複製元の記憶装置10, 20, 30の主論理ボリュームと複製先の記憶装置10, 20, 30の副論理ボリュームの内容を一致させるように行われる。

【0043】

また、リモートコピーの方式には同期方式と非同期方式とがあり、差分管理テーブル202の「複製方式」欄にて設定されている。同期方式の場合、複製元の記憶装置10, 20, 30は、情報処理装置11, 21から主論理ボリュームへのデータ書き込み要求を受信すると、まず自身の主論理ボリュームにデータを書き込む。また、複製元の記憶装置10, 20, 30は、書き込んだデータと同じデータを複製先の記憶装置10, 20, 30に送信する。複製先の記憶装置10, 20, 30は受信したデータを副論理ボリュームに書き込み、その旨を複製元の記憶装置10, 20, 30に通知する。そして、この通知を受信した複製元の記憶装置10, 20, 30は情報処理装置11, 21にデータの書き込みが完了した旨を通知する。

【0044】

このように同期方式では、主論理ボリュームと副論理ボリュームの双方にデータが書き込まれたことが確認された後に情報処理装置11, 21に完了通知が送信される。このた

め同期方式では情報処理装置 11, 21 が完了通知を受信した時点で主論理ボリュームの内容と副論理ボリュームの内容の一致性が確保されることになる。但し、副論理ボリュームへのデータの書き込みが完了するまでは情報処理装置 11, 21 に完了通知が報告されない。従って、同期方式では情報処理装置 11, 21 から複製元の記憶装置 10, 20, 30 にデータの書き込み要求が送信されてから、情報処理装置 11, 21 に完了通知が返ってくるまでのレスポンスタイムは一般に非同期方式の場合よりも長くなる。

【0045】

一方、非同期方式の場合、複製元の記憶装置 10, 20, 30 は、情報処理装置 11, 21 から主論理ボリュームへのデータ書き込み要求を受信すると、これに応じて主論理ボリュームにデータを書き込む。また、複製元の記憶装置 10, 20, 30 は、書き込んだデータと同じデータを複製先の記憶装置 10, 20, 30 に送信する。複製先の記憶装置 10, 20, 30 は受信したデータを副論理ボリュームに書き込み、その旨を複製元の記憶装置 10, 20, 30 に通知する。ここで複製元の記憶装置 10, 20, 30 は、主論理ボリュームにデータを書き込むと、複製先の記憶装置 10, 20, 30 の副論理ボリュームにデータが書き込まれたかどうかとは関係なく直ちに情報処理装置 11, 21 に完了通知を送信する。このため、非同期方式では一般に同期方式の場合に比べて情報処理装置 11, 21 へのレスポンスタイムが短くなる。しかしながら、主論理ボリュームと副論理ボリュームとの間のデータの一致性は必ずしも保証されない。なお、非同期方式の場合には、複製元の記憶装置 10, 20, 30 は、複製先の記憶装置 10, 20, 30 の副論理ボリュームに未反映のデータを管理している。

【0046】

本実施の形態においては、第一の主論理ボリュームと第三の論理ボリュームとのリモートペアは、同期方式でのリモートコピーを行っている。そのため、第一の主論理ボリュームと第三の論理ボリュームとは、常に一致した内容となっている。

【0047】

===差分管理テーブル===

差分管理テーブル 202 は、特定時刻を基準として、どの論理ボリューム 109 のどのブロックに対して更新が行われたかを示すものである。通常は図 5 に示すようにビットマップ形式で記憶装置 10, 20, 30 の共有メモリ 104 に記録されており、各ビットは論理ボリューム 109 内のブロックに対応づけられている。差分管理テーブル 202 は、初期状態は全て"0"の状態で記憶装置 10, 20, 30 の共有メモリ 104 に記録されている。記憶装置 10, 20, 30 は、論理ボリューム 109 へデータの書き込みを行うと、当該データが格納されているブロックを示す差分管理テーブル上のビットを"1"に更新する。

【0048】

本実施の形態においては、第一の論理ボリュームのローカルペアがスプリットされた時刻を基準として、第三の記憶装置 30 が第三の論理ボリュームのどのブロックを更新したかを、第三の記憶装置 30 の共有メモリ 104 に差分管理テーブル 202 として記録している。前述の通り、第一の主論理ボリュームと第三の論理ボリュームとのリモートペアは常に「ペア中」で、同期方式のリモートコピーが行われている。そのため、第三の記憶装置 30 の差分管理テーブル 202 は、第一の論理ボリュームのローカルペアがスプリットされた時刻以降に、第一の主論理ボリュームに書き込まれたデータの書き込み履歴を表している。つまり、差分管理テーブル 202 にて"1"がセットされているブロックの第三の論理ボリューム上のデータが、第一の論理ボリュームのローカルペアがスプリットされた時刻以降に、第一の主論理ボリュームに書き込まれたデータである。また、第一の情報処理装置 11 は、第一の論理ボリュームのローカルペアをスプリットした状態で、第一の副論理ボリュームを第二の論理ボリュームにバックアップしている。

【0049】

これにより、第一のサイトが被災した際には、第三の記憶装置 30 は、第三の記憶装置 30 の差分管理テーブル 202 上で"1"がセットされている第三の論理ボリュームのデー

タのみを、第二の論理ボリュームにリモートコピーすることで、第二の論理ボリュームを第一の主論理ボリュームと一致させることができる。

【0050】

ただし、第一の記憶装置10が第一の副論理ボリュームから第二の論理ボリュームへのリモートコピーを行っている最中に第一のサイトが被災した場合、第二の論理ボリュームは第一の論理ボリュームのローカルペアが前回にスプリットした時刻の第一の主論理ボリュームと一致した状態まで更新されていない。そのため、第三の記憶装置30は、第一の論理ボリュームが前回にスプリットした時刻からの差分管理テーブル202だけでは、第二の論理ボリュームを第一の主論理ボリュームと一致させることができない。

【0051】

そこで、本実施の形態においては、第三の記憶装置30は、第一の論理ボリュームのローカルペアが前回および前々回にスプリットした時刻からの差分管理テーブル202を記録しておく。これにより、第一の記憶装置11が第一の副論理ボリュームから第二の論理ボリュームへのリモートコピーを行っている最中に第一のサイトが被災した場合には、第三の記憶装置30は、第一の論理ボリュームのローカルペアが前々回にスプリットした時刻からの差分管理テーブル202を用いて第二の論理ボリュームを第一の主論理ボリュームと一致させることができる。

【0052】

第一の記憶装置10が第一の副論理ボリュームから第二の論理ボリュームへのリモートコピーを行っている最中に第一のサイトが被災した場合の例を、図7に示す。差分管理テーブル202には、第三の記憶装置30が更新を行った第三の論理ボリューム上のブロックについて、“1”を設定しているが、図7では説明の便宜上、論理ボリュームの書き込んだデータの内容を記している。図7に示す通り、被災した時刻には、第一の記憶装置10が行っている第一の副論理ボリュームから第二の論理ボリュームへのリモートコピーは、データ「5」までの更新が完了している状態である。その次にあるデータ「9」の更新が完了していないため、第二の論理ボリュームは第一の論理ボリュームのローカルペアが前回にスプリットした時刻の第一の主論理ボリュームと一致した状態とはなっていない。そのため、第三の記憶装置30は、第一の論理ボリュームのローカルペアが前々回にスプリットした時刻からの差分管理テーブルで“1”がセットされている、ブロック1からブロック6のデータ「1」、「5」、「9」、「3」、「4」、「2」を第二の論理ボリュームにリモートコピーする。これにより、第三の記憶装置30は、第二の論理ボリュームを第一のサイトが被災した時刻の第一の主論理ボリュームと一致させることができる。

【0053】

また、記憶装置10、20、30は、情報処理装置11、21から差分クリア要求を受信すると差分管理テーブル202のビットを全て“0”に更新する。情報処理装置11、21は、差分クリア要求を差分クリアの開始時刻を指定して行うこともできる。また、情報処理装置11、21は、差分管理テーブル202で更新管理を行っている論理ボリューム109がコンシステンシグループを形成している場合には、コンシステンシグループ単位で差分クリアを指示することができる。

【0054】

本実施の形態においては、前述の通り、差分管理テーブル202は第三の論理ボリュームごとに2つある。第三の記憶装置30は第一の情報処理装置11から差分クリア要求を受信すると、前々回に第一の論理ボリュームのローカルペアがスプリットした時刻からの差分管理テーブル202のみをクリアし、新たに今回スプリットした時刻からの更新情報を差分管理テーブル202に記録し始める。

【0055】

===通常運用時の遷移===

図8に、第一の情報処理装置11で本番系の運用を行っている際の、時刻T1から時刻T3の間における、各サイトのデータ状態の遷移を例示している。なお、論理ボリューム109中に記載しているSTnは、当該論理ボリューム109が時刻Tnの第一の主論理

ボリュームと一致していることを表している。また、 ΔT_{xy} は時刻 T_x から時刻 T_y の間の差分管理テーブル202を表している。

【0056】

まず、時刻 T_1 の状態を説明する。第一の論理ボリュームのローカルペアの前回スプリット時刻を T_0 とし、第一の情報処理装置11が時刻 T_1 に第一の論理ボリュームのローカルペアをスプリットした状態を示している。第一の論理ボリュームのローカルペアが「スプリット中」、第一の主論理ボリュームと第三の論理ボリュームのリモートペアが「ペア中」、第一の副論理ボリュームと第二の論理ボリュームのリモートペアが「スプリット中」の状態となっている。第一の情報処理装置11は、第一の論理ボリュームのスプリットと同時刻 T_1 を指定して、差分管理テーブル202のクリア要求を第三の記憶装置30に送信する。これを受信した第三の記憶装置30は、時刻 T_1 に、第一の論理ボリュームのローカルペアが前々回にスプリットした時刻からの差分管理テーブル202をクリアする。また、第二の論理ボリュームは第一の主論理ボリュームの T_0 時刻の状態となっている。

【0057】

次に、時刻 T_2 の状態を説明する。第一の記憶装置10が、第一の情報処理装置11からの指示により、第一の副論理ボリュームと第二の論理ボリュームとのローカルペアを再形成している状態を示している。第一の情報処理装置11は、第一の副論理ボリュームと第二の論理ボリュームのリモートペアのペア状態を監視しており、ペア状態が「ペア中」になると、当該リモートペアをスプリットするよう第一の記憶装置10に指示する。

【0058】

最後に、時刻 T_3 の状態を説明する。第一の情報処理装置11は、第一の記憶装置10に対して第一の論理ボリュームのローカルペアを再形成するよう指示する。これを受信した第一の記憶装置10は、第一の論理ボリュームのローカルペアを再形成する。

【0059】

このように、第一のサイトが被災していない間は、時刻 T_1 の状態から時刻 T_3 の状態の遷移を繰り返している。この遷移における第一の情報処理装置11および第三の記憶装置30の動作を説明する。

【0060】

まず、図9を用いて第一の情報処理装置11の動作を説明する。第一の情報処理装置11は、第一の記憶装置10に対して、第一の論理ボリュームのローカルペアを時刻指定でスプリットするよう指示する(S911)。ここで、スプリットはローカルペアに対してコンシステンシグループ単位で実行される。また、第一の情報処理装置11は、第三の記憶装置30に対して、前記スプリットと同時刻を差分クリア開始時刻として設定した差分クリア要求を送信する(S912)。なお、差分クリア要求は、スプリットが行われたローカルペアと同一のコンシステンシグループのリモートペアに対して行われる。設定した時刻の経過後(S913)、第一の情報処理装置11は、第一の記憶装置10からのスプリット完了通知と第三の記憶装置30からの差分クリア完了通知とを受信すると(S914)、第一の記憶装置10に対して、第一の副論理ボリュームと第二の論理ボリュームのリモートペアを再形成するよう指示する(S915)。第一の情報処理装置11は、第一の副論理ボリュームと第二の論理ボリュームのリモートペアのペア状態を監視し、当該ペアの状態が「ペア中」になると(S916)、第一の記憶装置に当該ペアのスプリットを指示する(S917)。次に、第一の情報処理装置11は、第一の論理ボリュームのローカルペアを再形成するよう第一の記憶装置10に指示する(S918)。第一の情報処理装置11では、この一連の処理を繰り返し実行している。

【0061】

次に、図10を用いて第三の記憶装置30の動作を説明する。第三の記憶装置30は、第一の情報処理装置11から差分クリア要求を受信すると、差分クリア開始時刻をコンシステンシグループ管理テーブル200に設定する(S1011)。第三の記憶装置30は、差分クリア開始時刻が設定されている間は、第一の記憶装置10からのデータの書き込み要

求を監視している (S1012)。第三の記憶装置 30 は、第一の記憶装置 10 からデータの書き込み要求を受信すると、当該データの書き込みが差分管理テーブル 202 をまだクリアしていない論理ボリューム 109 に対する書き込みかどうか確認する (S1013)。書き込みが差分管理テーブル 202 をまだクリアしていない論理ボリューム 109 に対するものである場合、第三の記憶装置 30 は、書き込み要求に設定された時刻と差分クリア開始時刻を比較する (S1014)。第三の記憶装置 30 は、書き込み要求に設定された時刻が差分クリア開始時刻よりも遅い場合は、先に、書き込み対象の論理ボリューム 109 の差分管理テーブル 202 をクリアする (S1015)。その後で、第三の記憶装置 30 は、書き込みを行った情報を差分管理テーブル 202 に記録する (S1016)。また、第一の記憶装置 10 からのデータの書き込み要求が、差分を既にクリアした論理ボリューム 109 に対するものである場合、第三の記憶装置 30 は、差分管理テーブル 202 のクリアを行わずに、書き込みを行った情報を差分管理テーブル 202 に記録する (S1017)。また、書き込みが差分管理テーブル 202 をクリアしていない論理ボリューム 109 に対するものである場合、書き込み要求に設定された時刻が差分クリア開始時刻よりも早い場合は、第三の記憶装置 30 は、差分管理テーブル 202 のクリアを行わずに、書き込みを行った情報を差分管理テーブル 202 に記録する (S1017)。

【0062】

このように、第三の記憶装置 30 は、第一の記憶装置 10 から受信した書き込み要求に設定された時刻と、第一の情報処理装置 11 から受信した差分クリア要求に設定された差分クリア開始時刻とを比較し、差分管理テーブル 202 への書き込みと差分管理テーブル 202 のクリアを時刻順に行っている。つまり、第一の記憶装置 10 が行う第一の主論理ボリュームへの書き込み及び第一の論理ボリュームのローカルペアのスプリットと、第三の記憶装置 30 が行う差分管理テーブル 202 への書き込み及び差分管理テーブル 202 のクリアとの順序が保たれている。これにより、第三の記憶装置 30 の差分管理テーブル 202 は、第一の論理ボリュームのローカルペアがスプリットされた時刻以降に、第一の主論理ボリュームに書き込まれたデータの書き込み履歴であることが保証される。

【0063】

また、第三の記憶装置 30 は、第一の記憶装置 10 からの書き込み要求を受信していない場合には、差分クリア開始時刻の経過後に (S1018)、まだクリアしていない論理ボリューム 109 の差分管理テーブル 202 をクリアする (S1019)。第三の記憶装置 30 は、差分管理テーブル 202 のクリアをコンシステンシグループ単位で行うため、コンシステンシグループ内の全てのペアの差分管理テーブル 202 のクリアが完了しているかどうかを確認する (S1020)。第三の記憶装置 30 は、コンシステンシグループ内の全てのペアの差分管理テーブル 202 のクリアが完了すると、コンシステンシグループ管理テーブルから差分クリア開始時刻を削除し (S1021)、第一の情報処理装置 11 に対して差分管理テーブル 202 のクリア完了通知を送信する (S1022)。なお、差分管理テーブル 202 は前述の通り各論理ボリューム 109 に対して 2 つ存在しているが、第三の記憶装置 30 は、第一の情報処理装置 11 から差分クリア要求を受信すると、前々回に第一の論理ボリュームのローカルペアがスプリットした時刻からの差分管理テーブル 202 のみをクリアする。

【0064】

=== 第一のサイト被災時の流れ ===

第一のサイトが被災した際に、第二のサイトを本番系として運用するために第二の論理ボリュームを第一の主論理ボリュームと一致させる動作を説明する。まず、第二の情報処理装置 21 が、第二の論理ボリュームと第一の副論理ボリュームとのリモートペアのペア状態の取得を第二の記憶装置 20 に指示する。この指示を受けた第二の記憶装置 20 は、第二の記憶装置 20 の共有メモリ 104 にあるペア管理テーブル 202 を参照し、その内容を第二の情報処理装置 21 に送信する。

【0065】

次に、第二の情報処理装置 21 は、第三の論理ボリュームを主論理ボリューム、第二の

論理ボリュームを副論理ボリュームとしたリモートペアを形成するよう第二の記憶装置 20 に指示する。また、第二の情報処理装置 21 は、第二の論理ボリュームと第一の副論理ボリュームとのリモートペアが「リシンク中」であったかどうかを、第二の記憶装置 20 に送信する。

【0066】

第二の記憶装置 20 は、第二の情報処理装置 21 から、第三の論理ボリュームと第二の論理ボリュームとのペア形成指示を受信すると、第二の記憶装置 20 のペア管理テーブル 202 を更新する。また、第二の記憶装置 20 は、第三の記憶装置 30 に対して、当該ペアの形成指示と、第二の情報処理装置 21 から受信した第二の論理ボリュームと第一の副論理ボリュームとのペア状態を送信する。第三の記憶装置 30 は、これを受信すると、第三の記憶装置 30 のペア管理テーブル 202 を更新する。第三の記憶装置 30 は、第二の記憶装置 20 より受信した第二の論理ボリュームと第一の副論理ボリュームとのペア状態に応じて、第三の論理ボリュームから第二の論理ボリュームに対してリモートコピーを行う。

【0067】

つまり、第二の論理ボリュームと第一の副論理ボリュームとのリモートペアの状態が「リシンク中」でなかった場合は、第三の記憶装置 30 は、第一の論理ボリュームのローカルペアが前回にスプリットした時刻からの差分管理テーブル 202 を参照する。また、第二の論理ボリュームと第一の副論理ボリュームとのリモートペアの状態が「リシンク中」であった場合は、第三の記憶装置 30 は、第一の論理ボリュームのローカルペアが前々回にスプリットした時刻からの差分管理テーブル 202 を参照する。第三の記憶装置 30 は、参照した差分管理テーブル 202 上で“1”がセットされている第三の論理ボリュームのブロックのみを、第二の論理ボリュームへリモートコピーする。

【0068】

図 11A は、第二の論理ボリュームと第一の副論理ボリュームのリモートペアが「リシンク中」でなかった場合の例を示している。第一の論理ボリュームのローカルペアが前々回にスプリットした時刻を T_0 、前回にスプリットした時刻を T_1 とし、第一のサイトが時刻 T_3 に被災した状況を表している。第二の論理ボリュームと第一の副論理ボリュームとのローカルペアのペア状態が「スプリット中」であるため、第三の記憶装置 30 は、時刻 T_1 からの差分管理テーブル 202 (ΔT_{13}) に記録されているブロックを第三の論理ボリュームから第二の論理ボリュームにリモートコピーすることで、第二の論理ボリュームを ST_3 とし、第一の主論理ボリュームと一致させることができる。

【0069】

次に、図 11B は、第二の論理ボリュームと第一の副論理ボリュームのリモートペアが「リシンク中」であった場合の例を示している。第一の論理ボリュームのローカルペアが前々回にスプリットした時刻を T_0 、前回にスプリットした時刻を T_1 とし、第一のサイトが時刻 T_2 に被災した状況を表している。第二の論理ボリュームと第一の副論理ボリュームとのローカルペアの状態が「リシンク中」であるため、第三の記憶装置 30 は、時刻 T_0 からの差分管理テーブル 202 (ΔT_{02}) に記録されているブロックを第三の論理ボリュームから第二の論理ボリュームにリモートコピーすることで、第二の論理ボリュームを ST_2 とし、第一の主論理ボリュームと一致させることができる。

【0070】

以上、本実施の形態について説明したが、上記実施例は本発明の理解を容易にするためのものであり、本発明を限定して解釈するためのものではない。本発明は、その趣旨を逸脱することなく、変更、改良され得ると共に、本発明にはその等価物も含まれる。

【図面の簡単な説明】

【0071】

【図 1】 本実施の形態に係る情報処理システムの概略構成を示す図である。

【図 2】 本実施の形態に係るディスクアレイ装置の構成を示す図である。

【図 3】 本実施の形態に係る情報処理装置の構成を示す図である。

【図 4】本実施の形態に係る記憶装置におけるチャネル制御部の構成を示す図である。

【図 5】本実施の形態に係る記憶装置における共有メモリに記憶されるテーブルを示す図である。

【図 6】本実施の形態に係るペア管理テーブルを示す図である。

【図 7】本実施の形態に係る差分管理テーブルを示す図である。

【図 8】本実施の形態に係る情報処理システムにおいて、第一のサイトで本番運用が行われている際の、各記憶装置の状態の遷移を示す図である。

【図 9】本実施の形態に係る情報処理システムにおいて、第一のサイトで本番運用が行われている際の、第一の情報処理装置における処理の流れを説明するフローチャートを示す図である。

【図 10】本実施の形態に係る情報処理システムにおいて、第一のサイトで本番運用が行われている際の、第三の記憶装置における処理の流れを説明するフローチャートを示す図である。

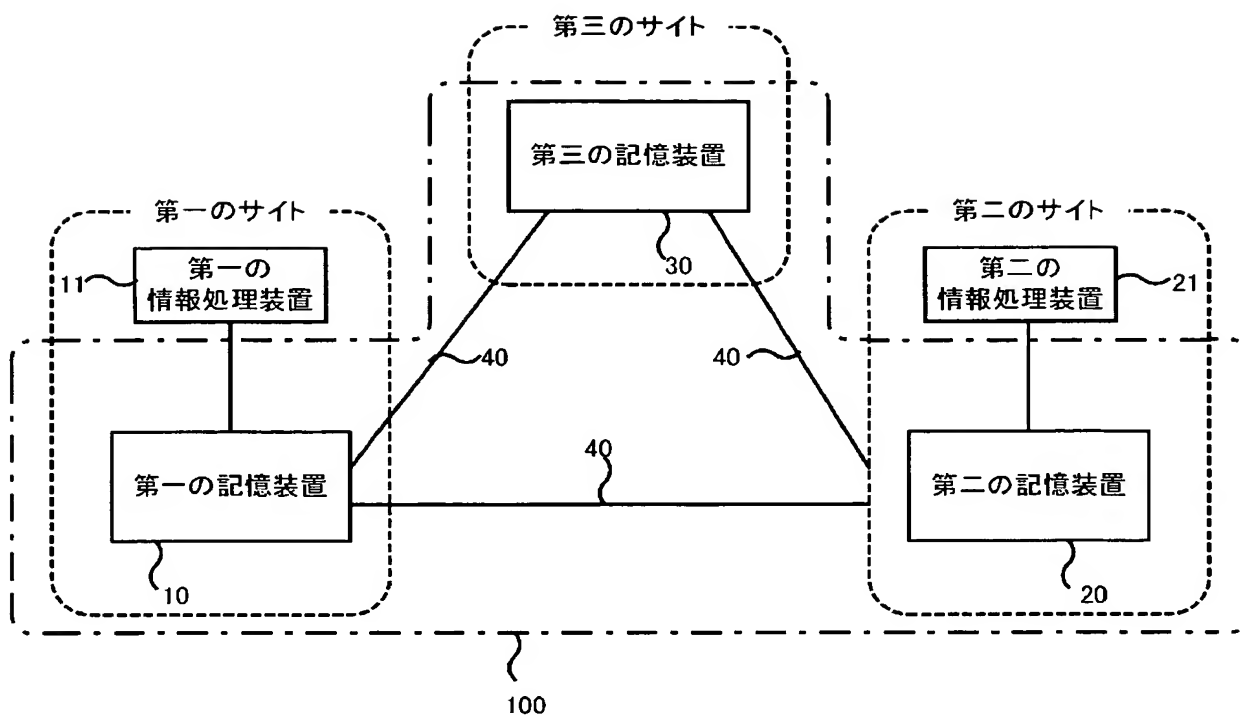
【図 11】本実施の形態に係る情報処理システムにおいて、第二のサイトを本番系に移行するために、第二の記憶装置の論理ボリュームを第一の記憶装置の論理ボリュームと一致させる動作を示す図である。

【符号の説明】

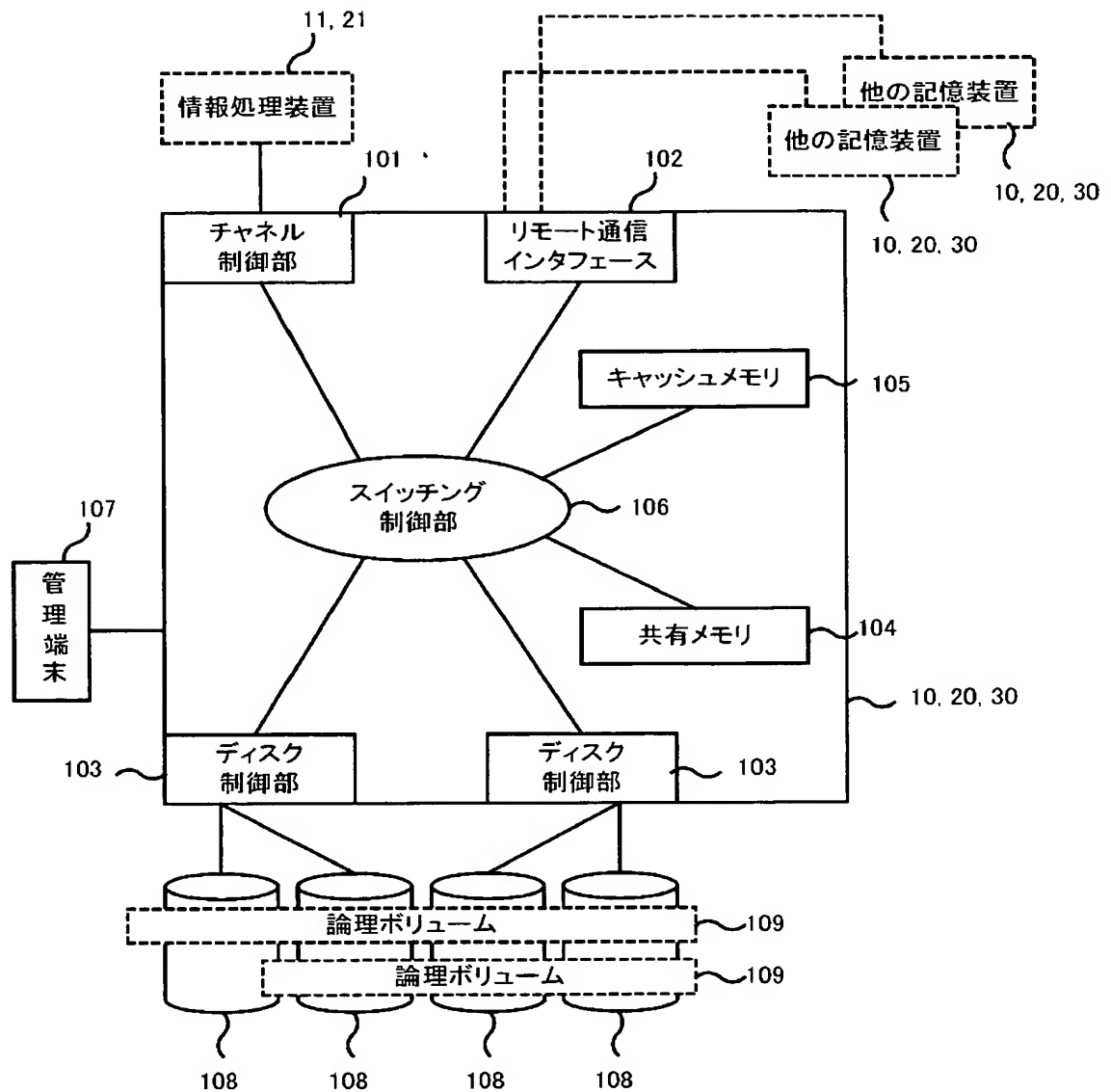
【0072】

10	第一の記憶装置	11	第一の情報処理装置
20	第二の記憶装置	21	第二の情報処理装置
30	第三の記憶装置	40	第一のネットワーク
100	ストレージシステム	101	チャネル制御部
102	リモート通信インタフェース	103	ディスク制御部
104	共有メモリ	105	キャッシュメモリ
106	スイッチング制御部	107	管理端末
108	記憶デバイス	109	論理ボリューム
110	CPU	120	メモリ
121	記憶デバイス管理プログラム	122	アプリケーションプログラム
130	ポート	140	記録媒体読取装置
150	入力装置	160	出力装置
170	記録媒体	200	コンシステンシグループ管理テーブル
201	ペア管理テーブル	202	差分管理テーブル
211	CPU	212	キャッシュメモリ
213	制御メモリ	214	制御プログラム
215	ポート	216	バス

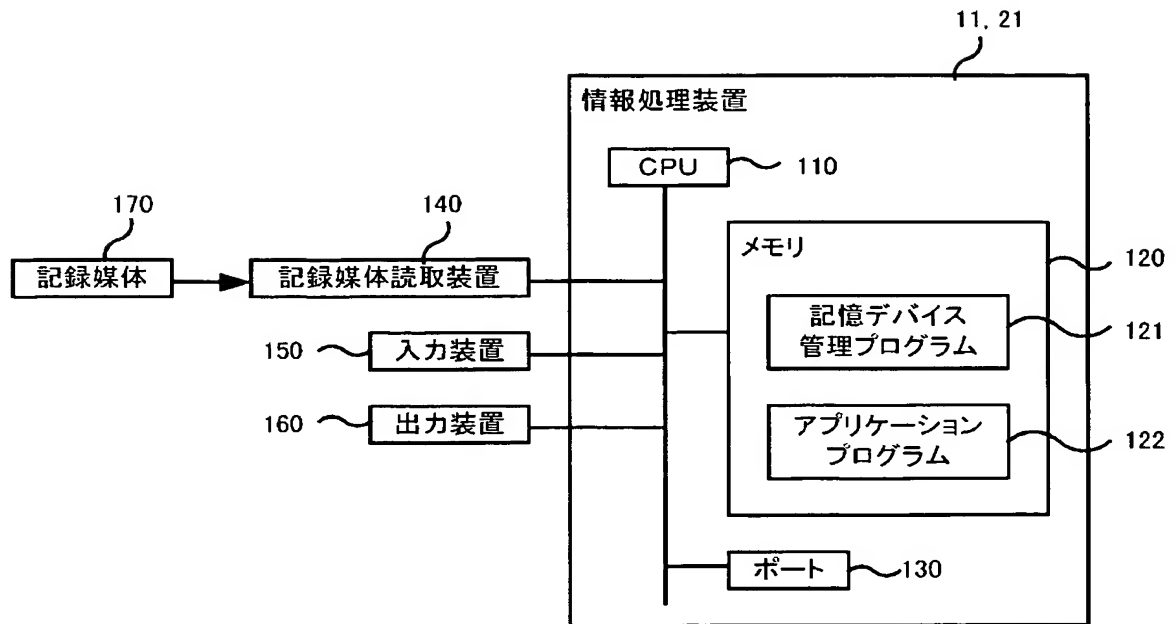
【書類名】 図面
【図 1】



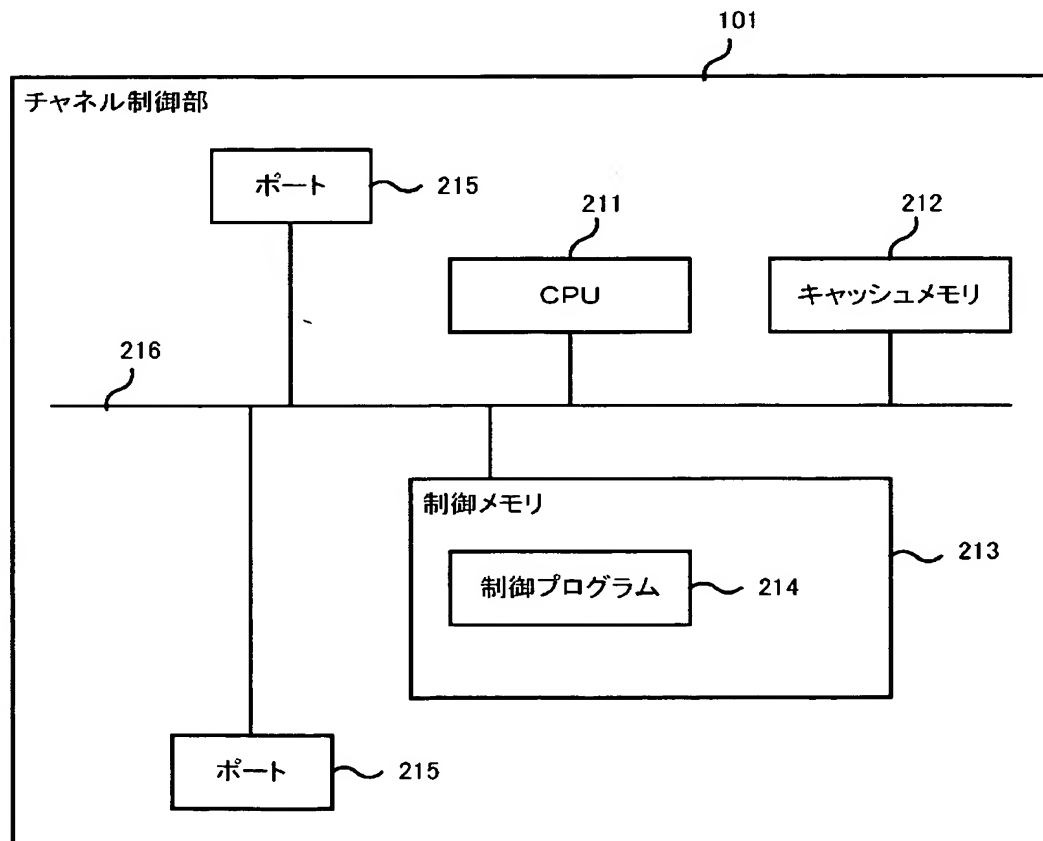
【図 2】



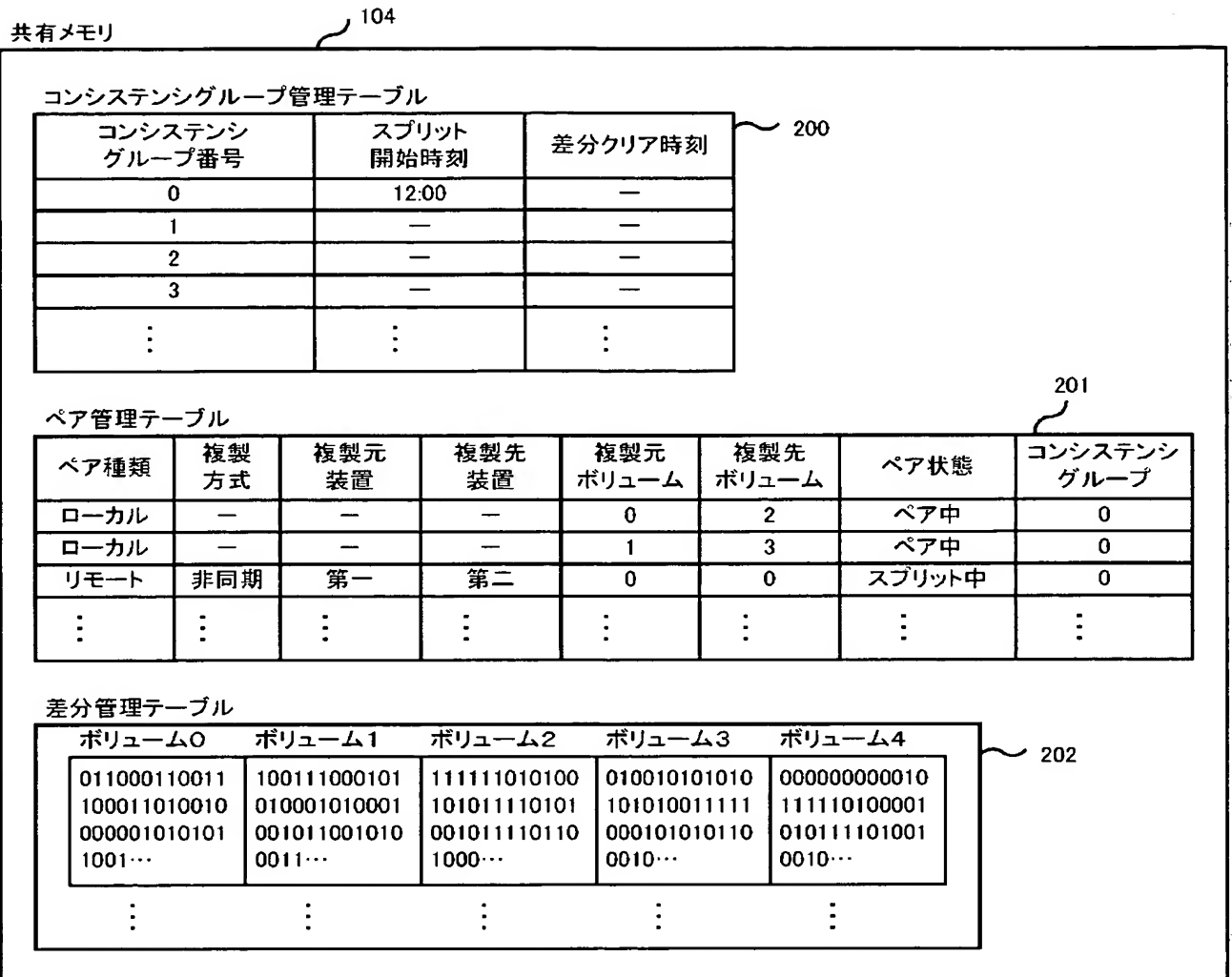
【図 3】



【図 4】



【図 5】



【図 6】

各記憶装置のペア管理テーブル

第一の記憶装置

201

ペア種類	複製方式	複製元装置	複製先装置	複製元ボリューム	複製先ボリューム	ペア状態	コンシステンシグループ
ローカル	—	—	—	0	2	ペア中	0
ローカル	—	—	—	1	3	ペア中	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
リモート	非同期	第一	第二	0	0	スプリット中	1
リモート	非同期	第一	第二	1	1	スプリット中	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
リモート	同期	第一	第三	0	0	ペア中	0
リモート	同期	第一	第三	1	1	ペア中	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

第二の記憶装置

201

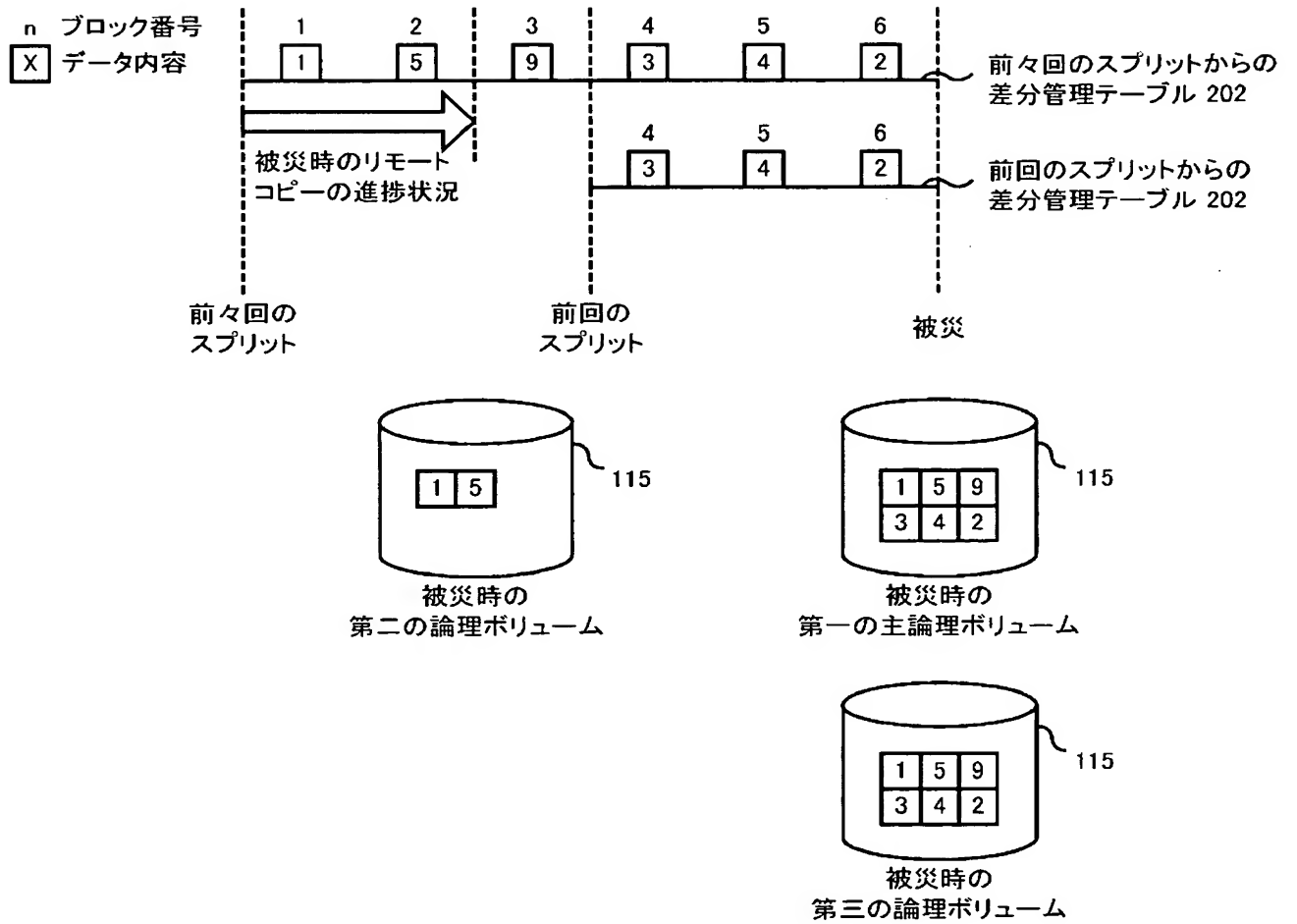
ペア種類	複製方式	複製元装置	複製先装置	複製元ボリューム	複製先ボリューム	ペア状態	コンシステンシグループ
リモート	非同期	第一	第二	0	0	スプリット中	1
リモート	非同期	第一	第二	1	1	スプリット中	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

第三の記憶装置

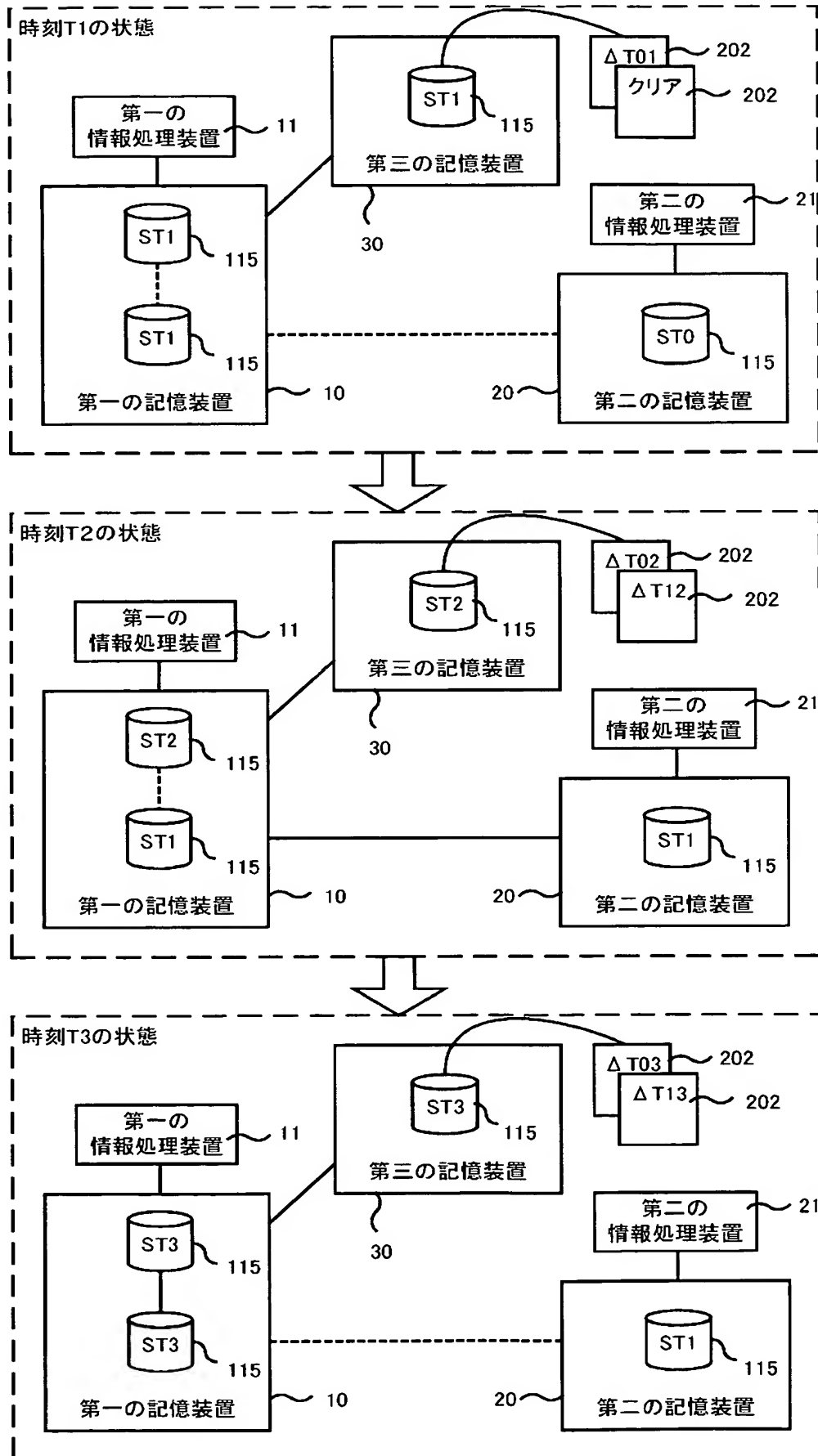
201

ペア種類	複製方式	複製元装置	複製先装置	複製元ボリューム	複製先ボリューム	ペア状態	コンシステンシグループ
リモート	同期	第一	第三	0	0	ペア中	0
リモート	同期	第一	第三	1	1	ペア中	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

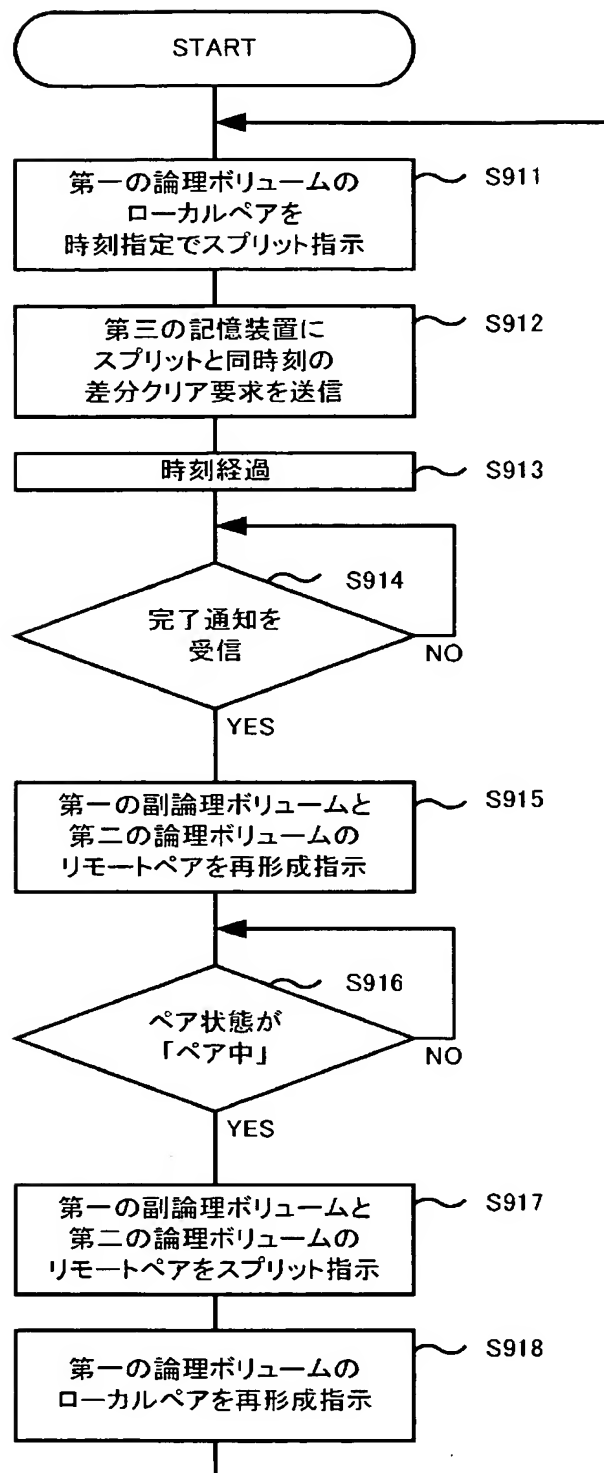
【図 7】



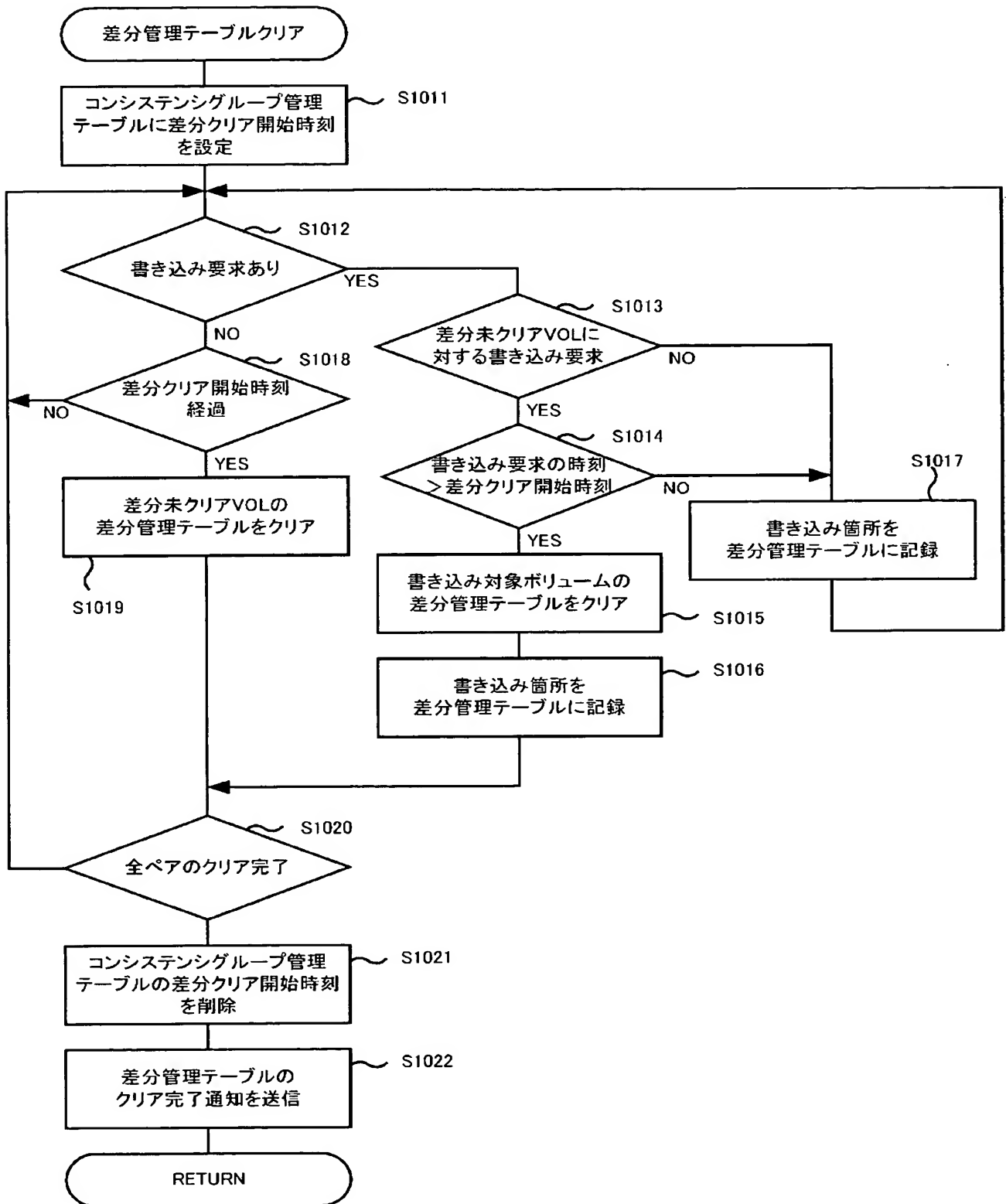
【図 8】



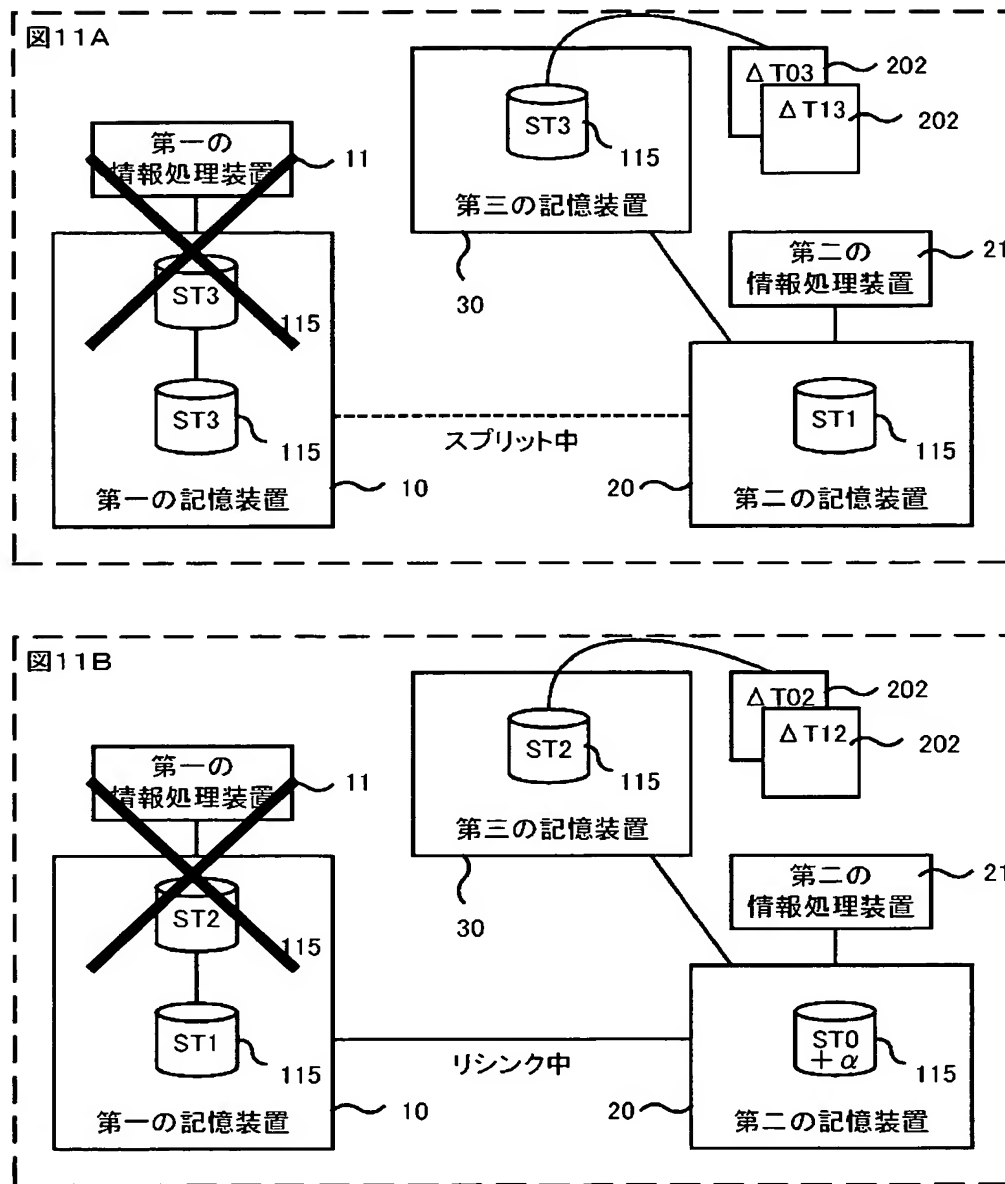
【図 9】



【図10】



【図 11】



【書類名】 要約書

【要約】

【課題】 第一のサイトが被災した際に、第二のサイトで迅速に処理を再開する。

【解決手段】 第一の記憶装置を含む第一のサイトと、第二の記憶装置を含む第二のサイトと、第三の記憶装置を含む第三のサイトとで構成される情報処理システムにおいて、第一のサイトが被災した場合に、第一の記憶装置の論理ボリュームと第二の記憶装置の論理ボリュームとの差分データのみを、第三の記憶装置の論理ボリュームから第二の記憶装置の論理ボリュームにリモートコピーすることで、第二の記憶装置の論理ボリュームを第一の記憶装置の論理ボリュームと一致させるステップを備える。

【選択図】 図 1

特願 2 0 0 3 - 3 0 9 1 9 4

出 願 人 履 歴 情 報

識別番号 [0 0 0 0 0 5 1 0 8]

1. 変更年月日	1 9 9 0 年 8 月 3 1 日
[変更理由]	新規登録
住 所	東京都千代田区神田駿河台 4 丁目 6 番地
氏 名	株式会社日立製作所